Curriculum Vitae   Gabor T. Marth, D.Sc.

Office: National Center for Biotechnology Information, NLM, NIH
Bldg. 45, Room 5AS29, 45 Center Drive, Bethesda, MD 20894
Tel: (301) 594-7087
Fax: (301) 480-0109
Email: marth@ncbi.nlm.nih.gov
Web site: http://www.ncbi.nlm.nih.gov/CBBResearch/Marth

Home: 6307 Crathie Lane
Bethesda, MD 20816
Tel: (301) 229-1163

Email: GaborMarth@yahoo.com

Education and Employment

Staff scientist, Computational Biology Branch, National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Maryland (2000-present).

Post-doctoral research associate, Genome Sequencing Center, Department of Genetics, Washington University School of Medicine, St. Louis, Missouri (1994-2000).

D.Sc. (Doctor of Science) degree in Systems Science and Mathematics, Washington University, St. Louis, Missouri (1988-1994).

Co-manager, Center for Robotics and Automation, Washington University (1992-1994).

Doctoral fellowship, Hungarian Academy of Science, Budapest, Hungary (1987-1988).

Internship, Institute for Heavy Ions Research, Darmstadt, Germany (1987).

B.S. and M.S. degrees in Electrical Engineering, Technical University of Budapest, Hungary (1983-1987).

Instructor, Cold Spring Harbor Lab Genome Informatics course (1999-present). Canadian Genetic Diseases Network Genome Informatics course (2000-present). Invited speaker, NIH Graduate course on Bioinformatics (2001). Engineering and Mathematics courses at the Washington University School of Engineering (1992-1993).

Member, Scientific Advisory Board, Marshfield Medical and Research Foundation Personalized Medicine Program, Marshfield, Wisconsin (2001-present).

Member, Theory Planning Group, proposed human haplotype map project (2001-present).

Interests

Quantitative approaches in Computational Biology. Genome-scale data mining and data interpretation. Study of fundamental evolutionary forces. Theoretical modeling and computer algorithmic development to describe the genome landscape of sequence variations. Dissection of the dynamics of population and species evolution through genealogy and mutation. Exploration of human origins. Understanding the genetic causes of heritable diseases. Integration of disease literature mining, sequence annotation, gene ontologies, and candidate gene approaches with haplotype marker information to enrich the experimental design for medical association studies.

Experience

Genome variation, population genetics, and pharmaco-genomics. Acquisition and analysis of genome-wide polymorphism and genotype data. Haplotype analysis and visualization tools. Modeling of genetic drift, population history, demography, mutation process, recombination, and selection for the analysis of human variations. Inferences on haplotype structure. Experimental design for medical association studies.

Single-nucleotide polymorphism (SNP) and deletion-insertion (DIP) detection. Primary developer of the PolyBayes SNP analysis method and software that incorporates multiple sequence alignment, algorithmic paralogue filtering and Bayesian-statistical SNP detection. SNP discovery in ESTs anchored to genomic clones. SNP detection in whole-genome shotgun sequence data, SNP detection in overlaps of large-insert genomic clones.

DNA sequencing software. Sequence assembly algorithmic design, data representations. Software for automated selection of finishing sequencing reactions. Hardware and software development for robotic sample re-arraying.

Publications

Journal articles and book chapters

1. Marth, G.T., Schuler, G., Yeh, R., Davenport, R., Agarwala, R., Church, D. et al. Sequence variations in the public human genome data reflect a bottlenecked population history. *PNAS* 100, 376-381. (2003).

2. Marth, G.T. Computational SNP discovery in DNA sequence data. *Methods Mol Biol* (Humana Press, ed. P.Y. Kwok) 212, 85-110. (2003).

3. Weber, J.L., David, D., Heil, J., Fan, Y., Zhao, C. & Marth, G.T. Human diallelic insertion/deletion polymorphisms. *American Journal of Human Genetics* 71, 854-62, (2002).

4. Vieux, E., Marth, G.T., & Kwok, P.Y. *SNP discovery and PCR-based assay design: From in silico data to the laboratory experiment.* In *Bioinformatics for Geneticists.* (ed. Barnes M. and Gray I.C., in press) (Wiley & Sons, 2002).

5. Marth, G., Yeh, R., Minton, M., Donaldson, R., Li, Q., Duan, S. et al. Single-nucleotide polymorphisms in the public domain: how useful are they? *Nature Genetics* 27, 371-2. (2001).

6. Sachidanandam, R., Weissman, D., Schmidt, S.C., Kakol, J.M., Stein, L.D., Marth, G. et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 409, 928-33. (2001).

7. Marth, G.T., Korf, I., Yandell, M.D., Yeh, R.T., Gu, Z., Zakeri, H. et al. A general approach to single-nucleotide polymorphism discovery. *Nature Genetics* 23, 452-6. (1999).

8. Dear, S., Durbin, R., Hillier, L., Marth, G., Thierry-Mieg, J. & Mott, R. Sequence assembly with CAFTOOLS. *Genome Research* 8, 260-7. (1998).

9. The Sanger Center and the Washington University Genome Sequencing Center. Toward a complete human genome sequence. *Genome Research* 8, 1097-108. (1998).

10. The C. elegans Sequencing Consortium. Genome sequence of the nematode C. elegans: a platform for investigating biology. *Science* 282, 2012-8. (1998).

Invited lectures

The structure of DNA sequence variation in the Human genome (National Institutes of Health, Bethesda, Dec. 2001)

Inferences on population history and linkage disequilibrium from genome-wide DNA variation data (4th International Meeting on Single Nucleotide Polymorphism and Complex Genome Analysis, Stockholm, Sweden, Oct. 2001).

DNA Polymorphism Data in Overlapping Genome Sequence (Marshfield Medical Foundation, Marshfield, Feb. 2001)

Recent presentations

Optimal marker sampling strategies for a general human haplotype map (5th International Meeting on Single Nucleotide Polymorphism and Complex Genome Analysis, Reykjavik, Iceland, Oct. 2002).

Models of genome-wide observed distributions of SNPs permit systematic description of human haplotype structure (Genome Meeting, Cold Spring Harbor, May 2002).

Inferences on human haplotype structure based on new population-genetic models of SNP density and allele frequency (Pacific Symposium on Bioinformatics, Kauai, Jan. 2002)

Unified analysis of SNP density and allele frequency spectra in genome-wide variation data (Genome Variation Meeting, Cold Spring Harbor, Sept. 2001).

SNPs in overlapping genomic sequences: Inferences on the genomic structure of variation and population history (Genome Meeting, Cold Spring Harbor, May 2001).